Geosensing Engineering and Mapping (GEM)
Civil and Coastal Engineering Department
University of Florida

Tutorial: Evaluating the Quality of ALSM Observations by Reading Artifacts in the Computed Surface Coordinates

GEM Center Report No. Rep_2005-01-002

B. Carter

January 11, 2005

Point of Contact:
Prof. K. C. Slatton
University of Florida; PO Box 116130; Gainesville, FL 32611
Tel: 352.392.0634, Fax: 352.392.0044, E-mail: slatton@ece.ufl.edu

Evaluating the Quality of ALSM Observations by Reading Artifacts in the Computed Surface Coordinates

Bill Carter
bcarter@ce.ufl.edu

Table of Contents

Evaluating the Quality of ALSM Observations by Reading Artifacts in the Computed Surface Coordinates

Bill Carter
bcarter@ce.ufl.edu

Introduction

ALSM is technologically complex and has far too many error sources to be reviewed in any detail in this brief overview. However, a number of the more common and largest sources of error leave “signatures” or “artifacts” in the processed data which make it possible for an analyst to detect and sometimes correct, even after the initial processing. Every ALSM data set will have such artifacts, if the data are examined at a sufficiently fine level. Below we discuss and show examples of the more common artifacts, using a commercially collected data set, that has significantly larger artifacts than is typically found in the research grade data produced by the NSF sponsored Center for Airborne Laser Mapping (NCALM). This data set was selected to help researchers who have limited or no experience in working with ALSM data to readily detect artifacts from a variety of sources.

Trajectory Errors

Error in the vertical component of the trajectory derived from the GPS and Inertial Measuring Unit (IMU) data is one of the most common errors in ALSM data. If the error is common to the entire survey area it generally can be traced to an error in the GPS base station, or an erroneous calibration parameter such as the antenna to sensor reference point vector. If the trajectory error suddenly changes during the collection of the data it is likely to be caused by a change in the geometry of the GPS constellation, such as the setting or loss of lock on one or more of the satellites. Sometimes trajectories will display anomalous slopes, which our experience suggests are generally caused by deficiencies in the GPS processing software. Some commercially available kinematic GPS software packages have difficulties handling tracks that are too far from the base station, often because they use single frequency GPS observations or do not have robust algorithms to do on-the-fly integer cycle fixing.

In the data set used here, there is one swath of data that is offset vertically from the otherwise relatively well matched swaths. The problem is immediately obvious on the contour map show in figure one, as a north-south running edge, cutting directly across the coverage. When a larger area is examined, another north-south edge can be seen where the vertically offset swath meets the adjacent swath on its other side. UF researchers quantified the vertical offset of this swath by comparing the ALSM heights to reference values derived from a kinematic GPS survey performed by the National Geodetic Survey (NGS). The offset is approximately 0.8 ft. The offset extends north and south of the airport area shown in Figure 1, but we have no ground-truth data beyond the airport area and can not say with any certainty if the vertical offset remains constant throughout the north-south extent of the swath. It is possible that the swath is tilted, which would cause the offset to grow larger or smaller beyond the airport area. Additional ground-truth values are needed to determine the extent and consistency of this problem. The offset between the two swaths is also clearly visible in the shaded relief image map show in Figure 2.

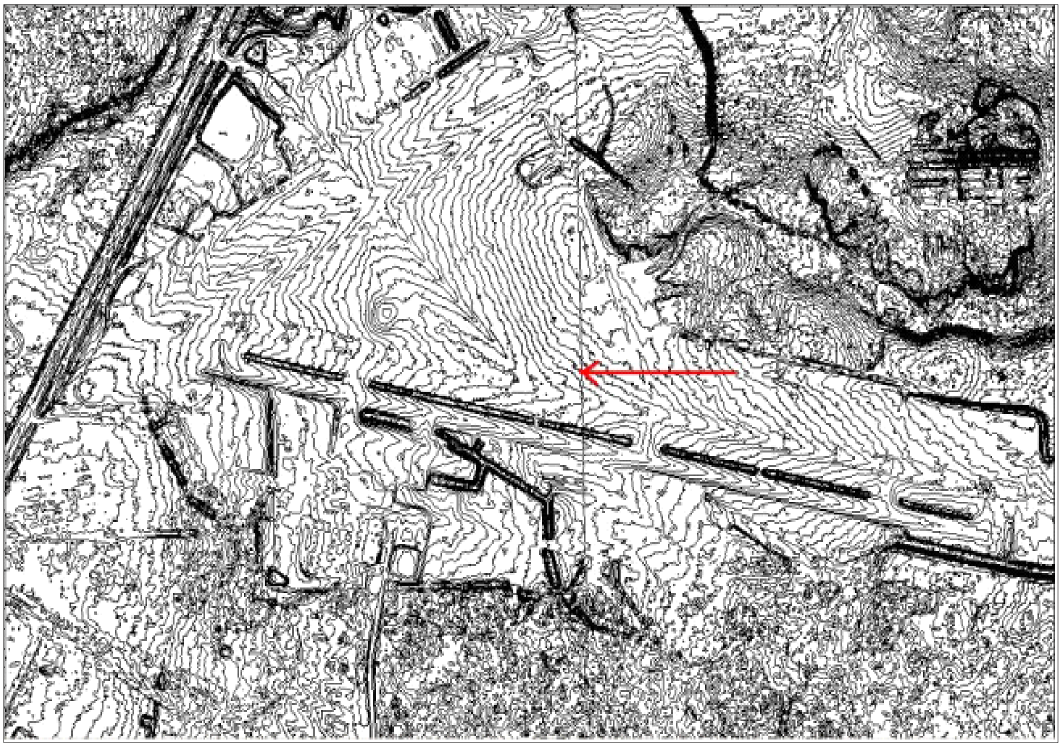

Figure 1. One-foot contour map. The red arrow indicates the location of a data artifact.

Figure 2. Shaded relief image map. The red arrow points to the artifact.

Data Voids

Data voids have two primary causes: failure to achieve contiguous coverage (gaps between adjacent swaths) during the data collection, and areas within the bounds of a swath where no return signals were recorded, i.e. areas with very low reflectivity, or off nadir areas with specular surfaces such as placid water. The former data gaps are often called “data holidays,” while the later are referred to as “dropouts.” We did not detect any holidays in the coverage in the test area. However, we did find significant numbers of dropouts (voids or sparse data) in areas with dark pavement and standing water. In particular, the apron areas around a number of the hangers in the airport complex have darker, lower reflective, asphalt surfaces that resulted in sparse or no returns. There are also areas painted black on the runways that resulted in data voids. Below is an example of dropouts on the dark asphalt surrounding hangers in the northwest quadrant of airport. The dimensions of the asphalt are approximately 350 feet by 390 feet, while the hangers are 330 feet by 50 feet.

Figure 3. Laser shots from raw file plotted over an area of dark asphalt surrounding hangers at the airport.

In this example there are fewer than 25 laser shots on 103,000 square feet of asphalt, and these occur almost exclusively on the paint stripes.If the dropouts were limited to the surfaces mentioned at the airport they would likely not be a significant problem. Unfortunately, the problem extended to road surfaces outside the airport complex,. There is no way to generate the missing data point from the ALSM observations, but it would be possible to survey the road surface in the missing areas and "patch" them into the ALSM coverage, or simply put them into a common data GIS data base to obtain a complete DEM. It is impossible to tell how time consuming and costly this process might be without doing a county-wide analysis of the extent of the data voids, and that is beyond the scope of this contract.

Figure 4. Laser shots from raw file posted on a section of a city street in another area of the survey.

In Figure 4 (above) contains 2000 linear feet of dark pavement on the travel lanes of a city street, which are devoid of laser coverage. The street surface was checked for dropouts for more than 4.7 miles, and we found that it contained only a handful of laser shots on the paved lanes, nearly all on the paint stripes.

Range Errors from Signal Level Variations

The strength of the return signals, commonly referred to as "intensity," typically varies by more than three orders of magnitude over an extended survey area. The response of the sensor, be it an avalanche photo diode (APD) or a photomultiplier tube (PMT), will not be perfectly linear over such a large dynamic range of signal. Manufacturers generally use constant fraction discriminators to minimize the "range walk" or systematic variation in range with signal level. In addition, some systems record the signal level and use a lookup table to correct the ranges.

The data set used here clearly displays artifacts from range walk, which are immediately visible in terrain height shaded relief images. Returns from the highly reflective painted portions of the runway yield short ranges, causing the surface height to be too high, by as much as about one foot (see Figure 7, following page). Figure 5 (below) is an ortho-photo of a portion of the airport. Note the locations of the paint striping. Figure 6 (following page) is a 3D surface image created from the filtered (ground) ALSM shots, with a vertical exaggeration factor of five. The red arrows point to range walk artifacts caused by a highly reflective surface (paint stripes). The blue arrow points out the vertically offset swath previously discussed.

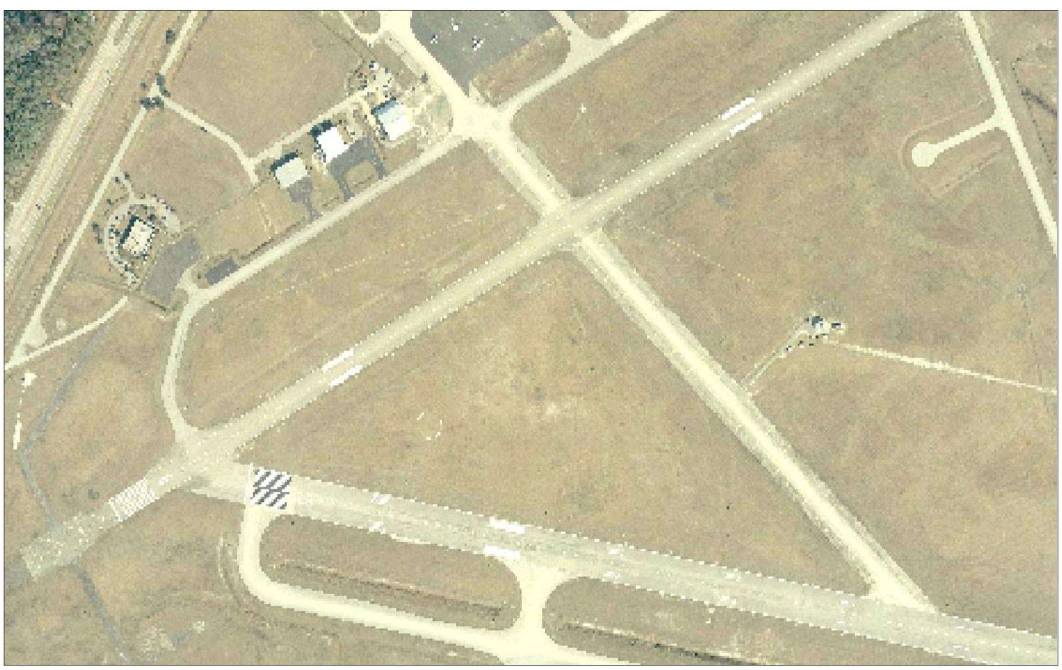

Figure 5. Aerial photograph showing paint stripes on the runways at GRA

Returns from low reflective surfaces, or from areas in forests where the returns may be from a portion of the footprint, yield long ranges, causing the surface height to be too low by as much as about two feet. If the data set included intensity values it would be possible to improve the calibration (lookup table) and then correct the surface heights to reduce the error from signal level variations. Unfortunately, this data set does not include intensity values and we know of no way to filter out the erroneous heights other than by tedious and time-consuming surface fitting. Even then, it is not possible to identify all those points affected by range walk, because large areas with many points may have significant errors if the reflectivity is near either extreme.

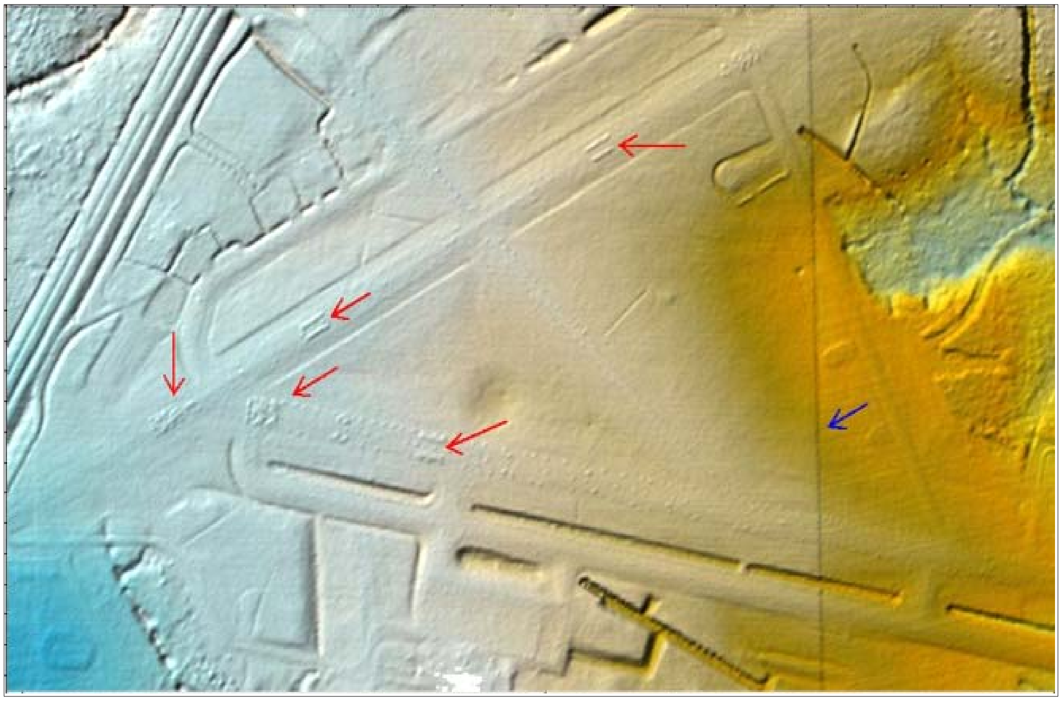

Figure 6. A 3D Surface Image showing range walk artifact.

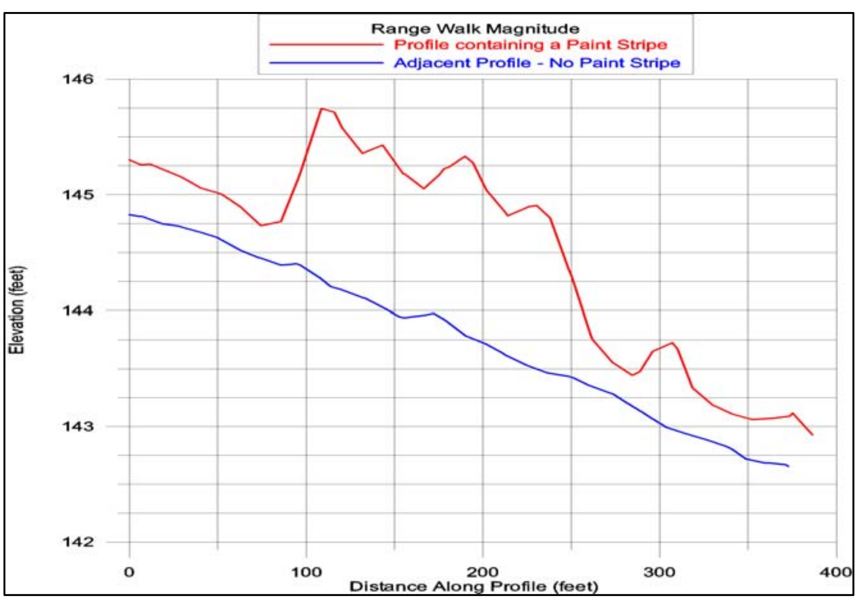

Figure 7. Elevation profiles illustrating the magnitude of the range walk error on paint striping at the GRA – up to 1 foot in height.

Variations in Data

Point Spacing

Most ALSM units operated from small single engine or dual engine aircraft are hard mounted - the sensor head is attached directly to the structural components of the aircraft, saving the cost, weight and complexities of a stabilized platform or gimbals. However, with no stabilized platform or gimbals to remove or compensate for changes in the orientation of the aircraft, rotations in the roll, pitch and yaw of the aircraft are mixed with those of the laser scanner, e.g., a single axis oscillating mirror laser scanner, effectively becomes a multi-axis scanning system.

An easy way to think about the effects of changes in aircraft orientation on the spacing of laser points on the ground is to picture the aircraft suspended above a flat surface, with zero roll, pitch, yaw and forward velocity. A simple oscillating mirror scanner would result in the laser points being distributed along a straight line perpendicular to the longitudinal axis of the aircraft, i.e., perpendicular to the nominal direction of flight. The line of laser pulses would extend equal distances on either side of the aircraft. The laser pulses would generally not hit identical points each scan, because the angular orientation of the mirror and the number of laser shots per unit time generally will not be integer multiples of one another, and there is no reason for the designer to invest in costly components to attempt to phase lock the laser shots and scan angle.

Consider what happens if the aircraft, still with zero forward velocity, is rotated about an axis running from wing tip to wing tip. For example, if the nose is rotated up, the tail rotates down, and the "pitch" of the aircraft is said to increase. It should be apparent that as the pitch increases, the laser scan line will be tilted forward and the laser points will hit the flat surface of terrain further toward the nose of the aircraft. For small changes in pitch angle the change in the "slant distance" to the ground will not increase dramatically, nor should the signal level drop significantly. But even a change in the pitch of one milliradian (approximately 3.5 arc minutes), when the aircraft is 1000 meters (3,200 feet) above ground level (agl) will move the points forward or aft, depending on the sign of the change in pitch, by 1.0 meter (3.3 feet). The displacement increases directly with height agl.

All aircraft, particularly small aircraft flying at relatively low altitudes, will continuously have some variation in the pitch angle. Changes of a few degrees over time scales of a few seconds, oscillating back and forth across the best "trim" conditions, are to be expected. As the pitch departs from the nominal vertical (in a positive or negative direction) the laser scan lines will grow progressively more widely spaced. As the scan line passes through nadir, the scan spacing will be minimum. With zero forward velocity the oscillations in pitch, combined with the laser scanner would cause the laser pulses to be distributed throughout a patch, extending forward and backward of the aircraft as well as to either side.

Now, consider what happens if the aircraft forward velocity is not zero, but a more typical value such as 60 meters per second. The forward motion of the aircraft is added to the forward and backward motion caused by the “scanning” associated with the ever-changing pitch. The result is that spacing between scan lines alternately becomes larger and then smaller, creating “bands” of denser and sparser spacing in the laser points on the surface of the terrain.

The data set used here was collected with a relatively high altitude agl. When the laser points are class posted, the expected banding is immediately obvious (see Figure 8 below). The void in the upper right is a dark asphalt road, and the circular void in the lower left is a water-filled depression. Note the different scan pattern of the adjacent swaths at either edge.

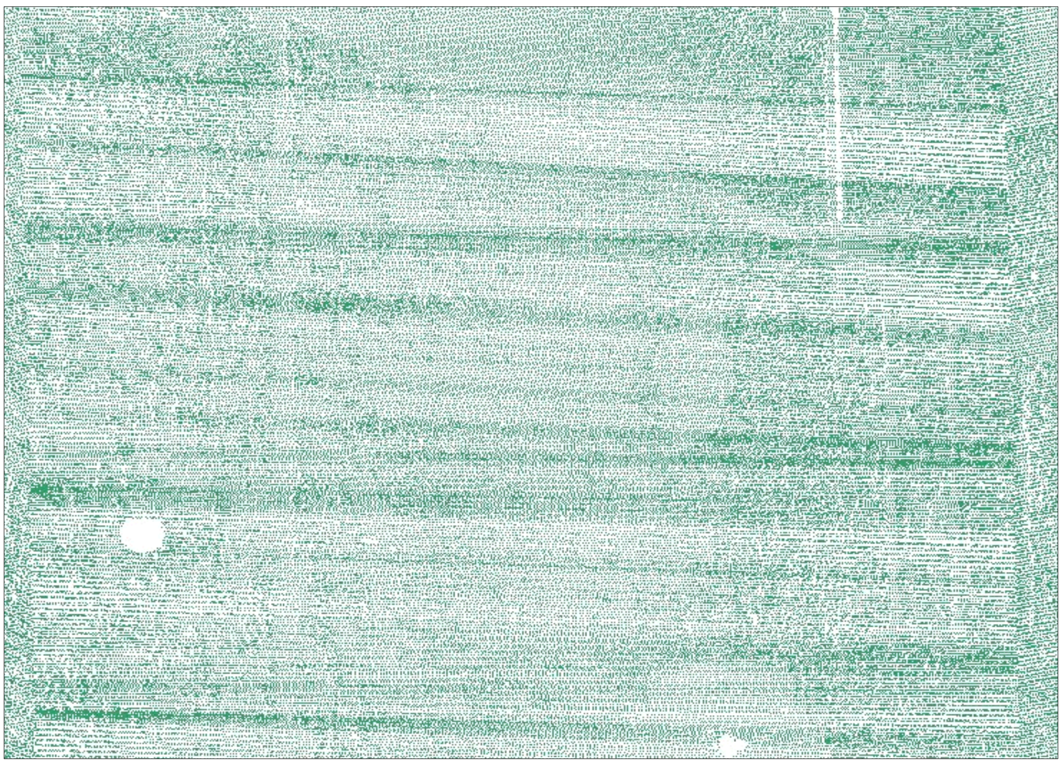

Figure 8. Plot of individual laser spots in a swath of ALSM, showing “banding” in the distribution of ground points.

The effect in the spacing of the data points is not in itself of much concern, because the data was collected with sufficient density that the variations do not result in excessive point spacing even with the maximum excursions in pitch. However, any error in the pitch calibration, or in the IMU pitch measurements, will result in errors in the final coordinates of the laser points on the surface, including the highly important height values. Qualitatively the presence of pitch error can be detected by the signature or artifacts they leave in the surface, i.e. undulations or waves in the surface with the peaks and troughs perpendicular to the flight path. In a shaded relief image, these undulations of the surface appear as alternating light and dark “bands” (wash boarding) with spatial scales of hundreds of meters, depending on the flying speed and the period of the oscillations in pitch.

Shaded relief images created from this data set are rich in the characteristic banding associated with height errors caused by the incomplete removal of the effects of variations in pitch of the aircraft. Figure 9 (below) is a shade, Near Gainesville, Florida. Note the banding in the open fields in the central part of the image, perpendicular to the North-South direction of flight. The purple areas indicate sparse information, such as the dark pavement of I-75, in the lower left.

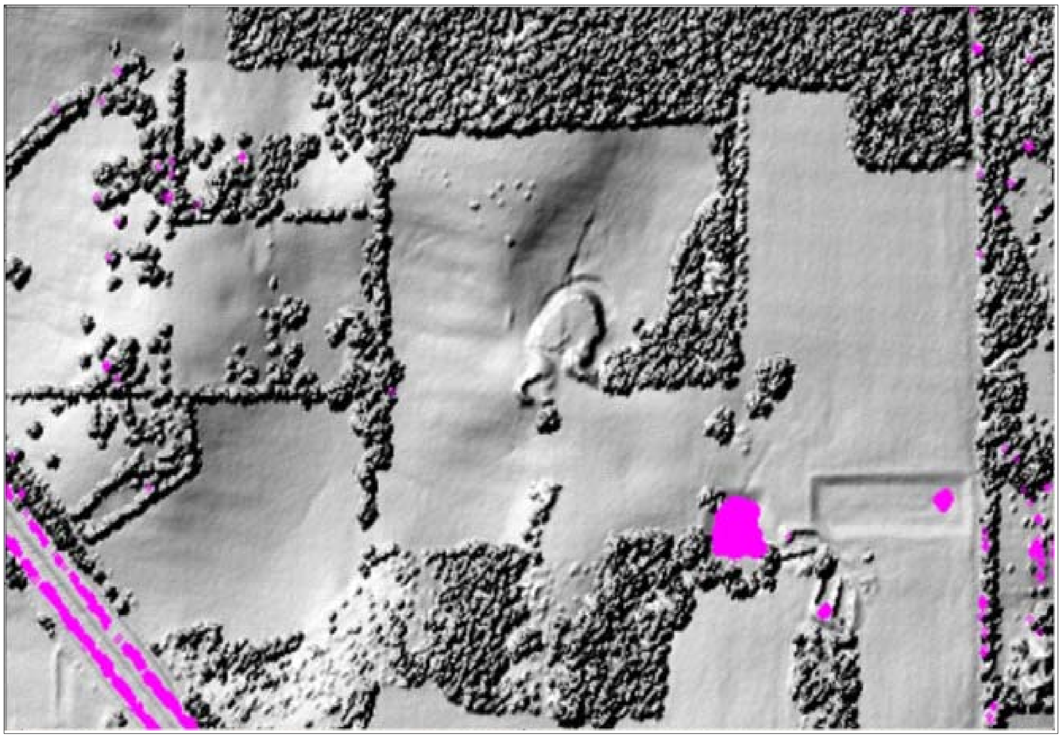

Figure 9. Shaded relief image which displays terrain height "banding" effects possibly caused by pitch error.

Quantifying the height errors caused by the residual pitch error is a much more difficult issue. Because the surface generally has natural undulations, albeit not generally oriented directly along the flight lines, it is problematical to slice a DEM and assume that all the undulations are artifacts caused by errors in the laser data. And because the magnitude and frequency of the variations in pitch generally change from flight to flight, the process of separating real surface undulations from artifacts can be very time consuming and the success can vary widely. The best approach to this difficult problem that we know of is based on analysis of data in the overlap of neighboring swaths. In this data set the overlap data were removed by the contractor before dlivering the data to the user. Because this is an important issue, we discuss the impacts of not having the overlap data in the next section.

Importance of Overlap Data

Because the cost of collecting ALSM data generally increases with increased overlap of neighboring swaths many operators tend to minimize the amount of planned overlap. In some cases the planned overlap may be so small that errors in aircraft flight line and orientation may actually result in data gaps or holidays. The data set used here appears to have been collected with sufficient overlap to avoid holidays.

Collecting overlapping data should not be considered as unproductive work, and simply an additional unjustified cost. To the contrary, the overlapping data is of great value to the analyst. For example, the apparent undulations of the surface caused by uncorrected pitch errors, discussed above, will generally vary swath to swath. By looking at apparent undulations in the overlap areas of neighboring swaths, or repeat swaths over the same area, it is possible to separate artifact undulations from real undulations of the terrain.

Unfortunately, because the data from the overlap areas have been removed during the processing of the data used here, it is much more difficult to isolate and quantify the errors in the laser point coordinates caused by aircraft orientation errors. We have focused our discussion of the aircraft orientation errors on pitch, but errors in roll and yaw (heading) cause similar errors, each causing their own distinctive signatures or artifacts in the data. For example, roll errors cause tilts across the swath, raising one edge and lowering the other. For a constant roll error the slope will remain constant and two adjacent swaths flown in the same direction show a step where they meet, while swaths flown in opposite directions will show a ridge or an inverted ridge along the line of intersection of the two tilted planes. Again, because of natural variations in the terrain it is not always immediately apparent that the swaths are tilted. But, the problem is much more apparent in overlap areas.

Based on our experience, UF researchers recommend that ALSM data be collected with substantial overlap – often as much as 50% overlap is justified because of the contribution it affords to quality controlling the data, and in removing residual problems from the data, after the initial standard processing. The data used here was collected with limited overlap and the data from the overlap were not included in the files delivered to the user. It is highly likely that the contractor has the data from the overlap areas and that it could be used to better quantify or perhaps even substantially reduce the errors in the final surface coordinates, but this would be a time consuming task.

Bare Earth DEM

The removal of ground clutter, such as points on the surfaces of buildings, cars trucks, utility poles and wires, and trees, to develop a “bare earth” digital elevation model is perhaps the most difficult facet of processing and interpreting ALSM observations. It is an area of ongoing research, and most providers of ALSM data are reluctant to share the algorithms and computer programs they have developed to perform this task. Because not all ALSM instruments observe or record the same information, algorithms developed to work with data from one system may not work well with data collected by another system. For example, some systems are capable of recording two or more returns per laser shot, while others may record only the first return.

Systems built by one major manufacturer record the first and last returns. Multiple returns, or first and last returns, can be used by an analyst to identify the presence of ground clutter, and to attempt to remove it. As mentioned above, if available, the intensity of the return signals can also assist in identifying and removing ground clutter.

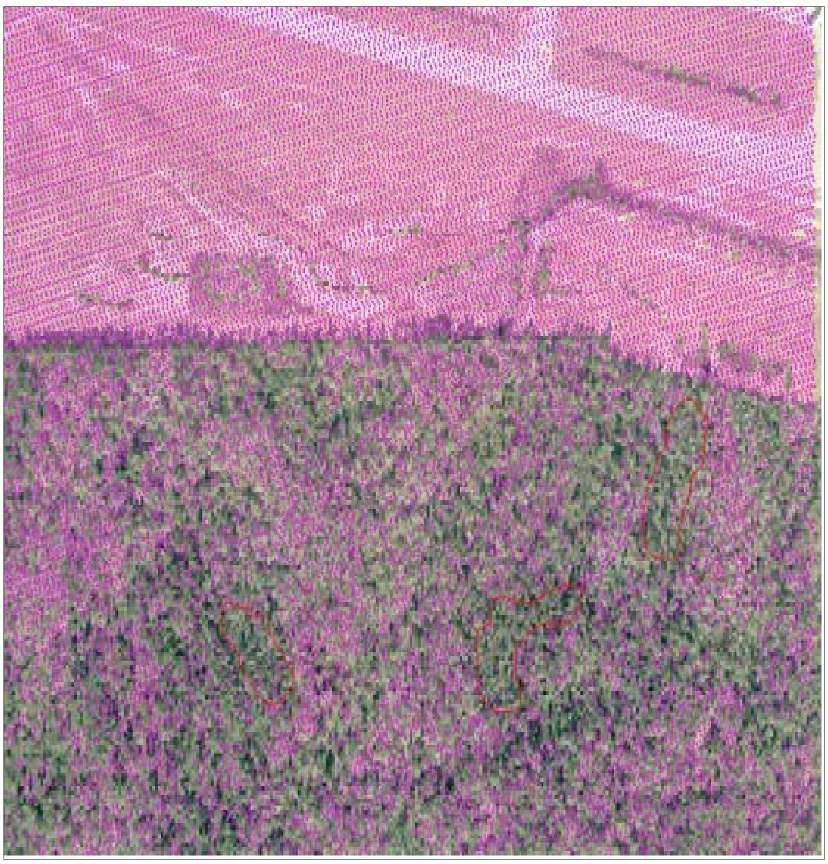

Figure 10. Ortho-photo of a portion of GRA showing boundar Between open and forested areas along with the laser shots that passed the vegetation filter.

It is likely that the system used to collect thedata set used here did record more than one stop per pulse, but UF researchers do not have files containing such data. The provider confirmed that the data set did not include intensity values. Our ability to quantify the quality of the bare earth DEM delivered by the contractor is severely limited by the lack of ground truth data in the areas of the airport complex covered by trees and dense brush. The NGS data set covers only paved and open grass covered terrain in the complex. However, even if we had ground points in the forested areas it would only enable us to evaluate the bare earth DEM in this local setting.

Qualitatively, based on our experience, we can say that the bare earth DEM in the airport area looks to be of relatively high quality in all but the most heavily forested areas, where there often is a mix of trees and thick underbrush. There are scattered areas where the attempt to remove the vegetation clearly resulted in sparse coverage by the remaining points (see Figure 10 below). In a shaded relief image of the filtered ALSM points (see Figure 11 below) these sparse areas result in overly smooth planar surfaces of various extent. Some of the larger areas devoid of laser shots in the filtered data set measure more than 100 feet across. There are also scattered "dimples" that indicate some remaining anomalously low elevations likely caused by unusually weak signals, as well as scattered "pimples" that indicate some remaining anomalously high elevations likely caused by failure to penetrate thick underbrush all the way to bare earth or unusually strong signals. Removal of such artifacts is not a perfect process and is generally tedious, time consuming and expensive. It may be a process that would be better left to be done in small areas on an as-needed basis. That is to say, areas should be carefully reviewed and interactively refined by an analyst when that portion of the bare earth DEM is to be used for a specific project or application.

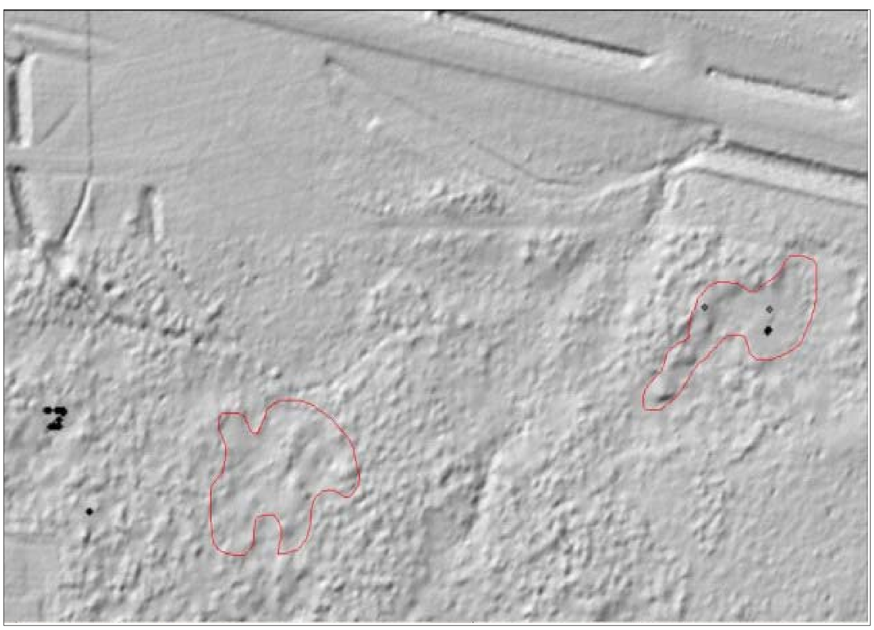

Figure 11. Shaded relief image of same area as above, showing effect of sparse coverage on surface roughness. Red circles show sparse coverage after filtering.

Concluding Remarks

ALSM units are complex optical-mechanical-electronic units that must be maintained and properly calibrated if the observations they collect are to be accurate. In addition, the position and orientation of the aircraft data obtained from GPS and IMU observations must be of high quality and processed using robust software. It is particularly important to have adequate overlap between adjacent (we recommend 50%) to enable the analyst to evaluate the consistency of the data quality and to confirm that there are not systematic errors caused by calibration problems, hardware malfunctions, or poor GPS solutions. With laser pulse rates of 25,000 to 100,000 pulses per second, and slowly varying deviations of such parameters as roll, pitch, yaw, and the X,Y,Z coordinates of the aircraft, adjacent range points have hightly correlated errors.

The artifacts displayed in the data used for here are signifcantly larger, by factors of 3 to 10, than the typical artifacts found in the research grade data produced by NCALM. Still, the NCALM data sets are not free of such artifacts and researchers should be aware of the existence of such artifacts, and be careful not to interpret them as real terrain features.